苦荞 *FtF5H* 基因克隆及表达分析

段迎[1]，杨晓琳[1]，蔡苏云[1]，贺润丽[1*]，尹桂芳[2]，王艳青[2]，卢文洁[2]，孙道旺[2]，王莉花[2]

（1. 山西中医药大学中药与食品工程学院， 太原 030619；2.云南省农业科学院生物技术与种质资源研究所/云南省农业生物技术重点实验室/农业部西南作物基因资源与种质创制重点实验室， 昆明 650201）

摘要：阿魏酸-5-羟基化酶（Ferulate 5-hydroxylase）是调控 S 型木质素合成的关键酶。为研究其在苦荞木质素生物合成途径中的分子机制，该文从苦荞转录组数据中筛选获得一个 F5H 基因，命名为 *FtF5H*(Genebank 登录号：MW455111)，采用生物信息学方法对苦荞 F5H 蛋白的理化性质、信号肽、跨膜结构、亚细胞定位、亲疏水性、蛋白质二级结构、蛋白质三级结构、氨基酸结构、系统进化树等进行分析和预测，并同时运用实时荧光定量PCR技术对 *FtF5H* 基因在厚果壳苦荞与薄果壳苦荞的叶、花、茎、果壳中的差异表达进行分析。结果表明：*FtF5H* 基因序列包含 1 395 bp 的完整 cDNA 开放阅读框，编码 464 个氨基酸；生物信息学预测 FtF5H 蛋白具有 P450 超家族结构，为亲水性稳定酸性蛋白，不具有跨膜结构域，且为非分泌性蛋白；FtF5H 蛋白的二级结构主要由α-螺旋和无规则卷曲组成；三级结构预测显示 FtF5H 蛋白与 5ylw.1.A 的相似度较高。系统进化分析表明 FtF5H 属于 CYP84A 亚家族。qRT-PCR 显示 *FtF5H* 基因在两种苦荞中的不同部位均有表达，且在厚果壳苦荞果壳中的表达量是薄果壳的 5 倍之高，表达具有极显著差异。该研究为进一步研究苦荞木质素合成的分子调控机制奠定了基础，对苦荞新品种的培育具有重要的意义。

关键词：苦荞，RT-PCR 克隆，阿魏酸-5-羟基化酶，生物信息学分析，基因表达

中图分类号：Q943

Cloning and expression analysis of *FtF5H* gene

收稿日期:2020-12-18

基金项目:国家自然科学基金地区科学基金项目(31460379)；国家燕麦荞麦产业技术体系"荞麦病虫害防控"(CARS-07-C-2)；山西省重点研发计划项目(201803D221012-6)；2019 年中医药公共卫生服务补助专项"全国中药资源普查项目(财社[2019]68 号)"[Supported by the National Natural Science Foundation of China (31460379); National Technical System of Oat and Buckwheat "Plant Diseases and Insect Pests Control of Buckwheat"(CARS-07-C-2); Key Research and Development Project of Shanxi (201803D221012-6); 2019 Subsidiary Special Project for Public Health Service of Traditional Chinese Medicine "Census of National Traditional Chinese Medicine Resource(Department of Social Security,Ministry of Finance [2019] No.68)"]

作者简介：段迎（1996-），硕士研究生，主要研究方向为中药资源开发与利用，（E-mail）18434376630@163.com。

*通信作者：贺润丽，女，博士，教授，研究方向为中药资源开发与利用，（E-mail）herunli666@163.com。

from tartary buckwheat

DUAN Ying[1], YANG Xiaolin[1], CAI Suyun[1], HE Runli[1*], YIN Guifang[2],

WANG Yanqing[2], LU Wenjie[2], SUN Daowang[2], WANG Lihua[2]

（1. *College of Traditional Chinese Medicine and Food Engineering*，*Shanxi University of Chinese Medicine*，Taiyuan 030619，China; 2. *Biotechnology and Genetic Germplasm Resources Research Institutes, Yunnan Academy of Agricultural Sciences/ Yunnan Provincial Key Laboratory of Agricultural Biotechnology/ Key Laboratory of Southwestern Crop Gene Resources and Germplasm Innovation, Ministry of Agriculture and Rural Affairs* , Kunming 650201，China）

Abstract： Ferulate 5-hydroxylase is a key enzyme that regulates the synthesis of S-type lignin.To study the function of ferulate 5-hydroxylase in lignin synthesis pathway of Fagopyrum tataricum, a *FtF5H*(Genebank accession number: MW455111) gene identified from Fagopyrum tataricum RNA-seq data was screened and analyzed in this research by bioinformatics methods including the physicochemical properties, signal peptide, transmembrane structure, subcellular localization, hydrophilicity, protein secondary structure and tertiary structure, amino acid structure, as well as phylogenetic tree. In addition, the real-time quantitative PCR was applied in this research to analyze the expression pattern of *FtF5H* gene in leaves, flowers, stems and husks of thick-husk tartary buckwheat and thin-husk tartary buckwheat. The bioinformatics analysis results were as follows: The length of the *FtF5H* gene was 1 395 bp, coding for 464 amino acids; the FtF5H protein was predicted to have'P450'superfamily structure. FtF5H protein, a non-secretory protein, is a hydrophilic, stable and acidic protein without transmembrane domain. The FtF5H protein secondary structure was predicted to be mainly composed of α -helix and random coil; the prediction of tertiary structure showed a high similarity with 5ylw.1.A. Phylogenetic analysis showed that FtF5H belonged to the CYP84A subfamily. In addition, qRT-PCR showed that the *FtF5H* gene was expressed in different parts of the two kinds of tartary buckwheat, and the expression level in the husk of the thick-husk tartary buckwheat was 5 times higher than that in the thin-husk,with extremely significant differences. The research establish a foundation for further study on the molecular regulation mechanism of lignin synthesis in Tartary buckwheat, and has important research significance for the cultivation of new tartary buckwheat varieties.

Keywords： tartary buckwheat，RT-PCR cloning，ferulate 5-hydroxylase，bioinformatics analysis，gene expression

苦荞(*Fagopyrum tataricum*)为药食两用作物，性味苦、平、寒，具益气力、续精神、利耳目、降气宽肠健胃等作用，被誉为“五谷之王” “三降”食品（朱云辉和郭元新，2014）。栽培苦荞通常果壳较厚，壳比率 20%~30%，果壳坚韧，脱壳率从 2%到 6%不等（Song et al.，2019），难以脱壳生产整粒苦荞米，不易制得高芦丁含量麸皮层全营养苦荞米，一定程度上降低了苦荞的活性物质与营养功效（陈庆富等，2015）；同时在苦荞种子直接磨粉加工中会有少量壳粉混入，导致适口性下降，也限制了苦荞产品的精深加工，严重影响苦荞产业的发展。‘小米荞’是我国云南、贵州等地的一种易脱壳地方品种，脱壳率最高达 93%（Song et al.，2019），作为一种特殊资源，‘小米荞’对研究苦荞薄果壳特性的遗传规律、功能基因、品种培育及加工生产等方面均具有重要的理论和现实意义。

F5H 属于细胞色素 P450 的单氧化物酶（Meyer et al.，1998），Meyer 于 1996 年第一次从香枫中分离得到，是 CYP84 家族成员，具有重要的生物学功能，该基因能催化阿魏酸、

松柏醇、松柏醛生成 5-羟基阿魏酸、5-羟基松柏醇和 5-羟基松柏醛，是调控 S 型木质素合成的关键酶（Humphreys & Chapple，2002）。目前，已经从拟南芥（*Arabidopsis thaliana*）（Franke et al.，2000）、当归（*Angelica sinensis*）（温随超，2015）、毛白杨（*Populus tomentosa*）（陈雪等，2015）、亚麻（*Linum usitatissimum*）（王进，2009）、油菜（*Brassica campestris*）（李扬等，2013）等植物中克隆到 F5H 基因，尚未发现 F5H 基因在苦荞中被克隆的相关报道。

吴朝昕（2020）对厚果壳和薄果壳苦荞的果壳成分研究发现：不同品种厚果壳苦荞的果壳木质素平均含量均高于薄果壳苦荞，课题组前期在‘云荞 1 号’和‘小米荞’杂交 F_2 代群体中，对厚果壳和薄果壳进行转录组测序比较分析（数据未发表），发现木质素生物合成途径大部分基因在苦荞厚果壳的表达量高于薄果壳，推测苦荞果壳厚可能由于积累了较多的木质素。转录组测序表明 F5H 基因在厚果壳的表达量显著高于薄果壳，为了研究 F5H 基因在苦荞木质素合成中的作用，本研究采用 RT-PCR 法克隆‘云荞 1 号’和‘小米荞’的 *FtF5H* 基因，对其进行生物信息学分析、实时荧光定量 PCR 验证，为进一步研究该基因功能提供依据。

1 材料与方法

1.1 材料

苦荞厚果壳材料‘云荞 1 号’为云南省农业科学院生物技术与种质资源研究所选育品种，薄果壳材料‘小米荞’为云南省地方品种。剥离不同发育期‘云荞 1 号’‘小米荞’的果壳，将剥离的果壳分别混合在一起作为基因克隆的材料；荧光定量 PCR 分析用‘云荞 1 号’、‘小米荞’不同组织部位材料：结实期的叶、花、茎、果壳。

1.2 RNA 提取与 cDNA 第一链合成

根据 Trizol 提取试剂盒说明进行总 RNA 提取，1.5%琼脂糖凝胶电泳检测所有样品 RNA 的完整性。cDNA 第一链合成：在 0.2 mL PCR 管中加入 5 μL 总 RNA、1 μL 随机引物、1 μL ddH_2O，70 ℃温浴 5 min。冰浴 2 min；离心加入 2.0 μL 5 × First-Strand Buffer、0.5 μL 10 mmol•L^{-1} dNTP、0.25 μL Rnase inhibitor、0.25 μL Reverse Transcriptase，总体系 10.0 μL，42 ℃温浴 60 min，72 ℃温浴 10 min。

1.3 *FtF5H* 基因克隆

根据转录组测序获得的阿魏酸-5-羟基化酶（*FtF5H*）基因的核苷酸序列设计 4 条特异引物（表 1）。RT-PCR 反应体系如下：cDNA 模板 1 μL、dNTPs（10 mmol·L^{-1}）0.2 μL、2×GC Buffer Ⅰ 12.5 μL、TaqDNA 聚合酶（5U·μL^{-1}）0.2 μL、ddH_2O 10.1 μL、F (10 μmol·L^{-1}) 0.5 μL、

R (10 μmol·L^{-1}) 0.5 μL。PCR 反应条件为 95 ℃预变性 3 min；94 ℃变性 30 s、58 ℃退火 30 s、72 ℃延伸 90 s、共循环 33 次；72 ℃修复延伸 7 min。PCR 产物用 1%琼脂糖凝胶电泳检测，把目的条带切胶回收，送生工生物工程（上海）股份有限公司测序。

表 1 引物序列及用途

Table 1 Sequence and usage of primers

用途 Usage	引物名称 Primer name	引物序列 Primer sequence（5'-3'）
RT-PCR	FtF5H-119F	TTCTTAGCCACGGCGGTCT
	FtF5H-1494R	GGTAACTCCCACGTGAAGCATTG
	FtF5H-1298F	AAGGACTCGTGGGATGACCCT
	FtF5H-1882R	ACATTGCATGCTTGATTGCAGTATA
荧光定量 PCR qRT-PCR	H3-F	AAGAAGTCCCACAGATACCGC
	H3-R	AGCCTCCTGAAGAGCTAGCAC
	FtF5H-F	TCGGTTCGTGATGAGGTGG
	FtF5H-R	CCCGGTAGATTACGTTCTTTGT

1.4 克隆基因的生物信息学分析

利用 NCBI ORF finder 在线程序及 Conserved domains 数据库对测序获得的 cDNA 序列进行开放阅读框及保守功能结构域分析；通过 ProtParam 对蛋白理化性质进行分析；利用 ProtScale 对蛋白亲疏水性进行分析；采用 TMHMM Server v. 2.0 分析编码氨基酸的跨膜结构域；SignalP 5.0 Server 预测信号肽；NetPhos 3.1Server 进行磷酸化位点预测；Psort 分析亚细胞定位情况。通过 SOPMA 及 SWISS-MODEL 预测蛋白的二级、三级结构；编码蛋白多重序列比对采用 DNAMAN 软件；Neighbor-joining 系统进化树用 MEGA6.0 构建（陈媞颖等，2017）。

1.5 实时荧光定量 PCR（qRT-PCR）分析

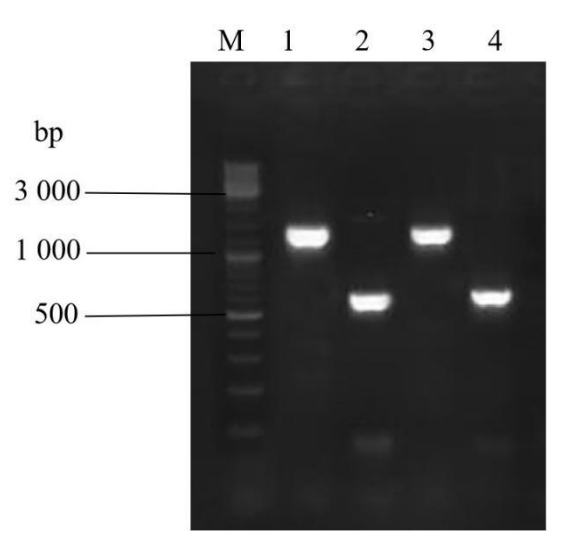
采用 Trizol 总 RNA 抽提试剂盒分别提取‘云荞 1 号’和‘小米荞’叶、花、茎、果壳 8 个组织的总 RNA，反转录得到第一链 cDNA，以其为模板进行 qRT-PCR 分析，每个样品设 3 次技术重复。荧光定量 PCR 反应体系：10 μL 2 × SG Fast qPCR Master Mix、2 μL cDNA 模板、上下引物各 0.4 μL、ddH_2O 7.2 μL。扩增程序：95 ℃ 3 min ，95 ℃ 5 s，60 ℃ 30 s，

循环 45 次。以苦荞 H3（JF769134.1）为内参基因，采用 $2^{-\Delta\Delta Ct}$ 法计算 *FtF5H* 的相对表达量，反应所用引物见表 1。

2 结果与分析

2.1 苦荞总 RNA 提取与 *FtF5H* 基因克隆

提取的总 RNA 经 1.5％琼脂糖凝胶电泳，在紫外透射光下观察 RNA 条带清晰可见，亮度比在 1∶1 至 2∶1 之间，表明 RNA 基本未降解，可用于后续基因克隆。利用 RT-PCR 分别克隆获得‘云荞 1 号’和‘小米荞’的 F5H 基因，产物经 1.0%琼脂糖凝胶电泳分析，电泳结果见图 1。对 RT-PCR 产物进行测序拼接，结果发现：‘云荞 1 号’和‘小米荞’的 F5H 基因核酸序列完全一致，且与转录组测序序列相同。ORF finder 分析结果表明，2 条基因序列均含有一个完整的开放阅读框，开放阅读框长度均为 1 395 bp，编码 464 个氨基酸，将该基因命名为 *FtF5H*（GenBank 登录号：MW455111）。利用 DNAMAN 将核苷酸序列翻译成氨基酸如图 2 所示。



M. SM0331 DNA Marker；1，2.‘云荞 1 号’*FtF5H* 基因克隆片段；3，4.‘小米荞’*FtF5H* 基因克隆片段。

M. M0331 DNA Marker; 1,2. *FtF5H* gene clone fragment of Yunqiao 1;

3，4. *FtF5H* gene clone fragment of rice buckwheat.

图 1　*FtF5H* 克隆片段的 PCR 电泳图

Fig. 1　The electrophoresis results of *FtF5H* fragments

```
    TGAAC
    ATGATGGACCAACTCACTCATCGTGGGTTGGCCCGTCTCGCCGGAATATACGGTGGAATATTCCATATGCGTATG
1    M  M  D  Q  L  T  H  R  G  L  A  R  L  A  G  I  Y  G  G  I  F  H  M  R  M
    GGTTTCCTCCATATGGTTACAATATCAAGCCCGGAAGCGGCCCGACAAGTCCTTCAAGTCCAAGACAACATCTTC
26   G  F  L  H  M  V  T  I  S  S  P  E  A  A  R  Q  V  L  Q  V  Q  D  N  I  F
    TCCAACCGGCCCGCTACCATCGCGATTAAATACCTTACTTACGATCGGGCCGACATGGCGTTCGCTCATTACGGG
51   S  N  R  P  A  T  I  A  I  K  Y  L  T  Y  D  R  A  D  M  A  F  A  H  Y  G
    CCTTTCTGGCGTCAAATGCGGAAACTCTCCGTCATGAAGTTATTCAGCCGTAAACGGGCCGAGTCTTGGCAATCG
76   P  F  W  R  Q  M  R  K  L  S  V  M  K  L  F  S  R  K  R  A  E  S  W  Q  S
    GTTCGTGATGAGGTGGATCGGATTCTCCGGGTCATGGATAAAGCTAACGGGTCACCCGTTAATATCGGCGAGCTT
101  V  R  D  E  V  D  R  I  L  R  V  M  D  K  A  N  G  S  P  V  N  I  G  E  L
    GTATTCTCTTTAACAAAGAACGTAATCTACCGGGCCGCGTTCGGGTCTAACTCGAACGAGGGACAAGATGAGTTC
126  V  F  S  L  T  K  N  V  I  Y  R  A  A  F  G  S  N  S  N  E  G  Q  D  E  F
    ATTTCTATATTGCAGGAGTTTTCGAAGTTGTTTGGAGCATTCAATTGGGCTGATTTTATTCCGTTTATCGGATGG
151  I  S  I  L  Q  E  F  S  K  L  F  G  A  F  N  W  A  D  F  I  P  F  I  G  W
    GCTGACCCGAACGGGCTTAATGCAAGGCTCGAGAAGGCCCGAAAATCTCTTGATAGGTTCATAGATGATATCATA
176  A  D  P  N  G  L  N  A  R  L  E  K  A  R  K  S  L  D  R  F  I  D  D  I  I
    AACGATCACTTGGAGAAGAAGAAGACGGGGATAAAGGGCTTGGAGGATGACATGGTAGATGAGTTGCTAGTGTTT
201  N  D  H  L  E  K  K  K  T  G  I  K  G  L  E  D  D  M  V  D  E  L  L  V  F
    TACAGTGAAGGTGACGGTAAAGTGACGGAGTCGGACGATTTACAGAACGCGATTCGACTCACTAAGGATAATATC
226  Y  S  E  G  D  G  K  V  T  E  S  D  D  L  Q  N  A  I  R  L  T  K  D  N  I
    AAGGCAATAATCATGGATGTGATGTTTGGCGGGACTGAAACTGTTGCATCGGCAATAGAATGGGCCATGTCGGAG
251  K  A  I  I  M  D  V  M  F  G  G  T  E  T  V  A  S  A  I  E  W  A  M  S  E
    CTAATGCATAGCCCGGAAGCCATGAAACGAGCTCAAAACGAGCTCGCCGAGGCAATCGGGTTTAACCGACCCGTG
276  L  M  H  S  P  E  A  M  K  R  A  Q  N  E  L  A  E  A  I  G  F  N  R  P  V
    GAAGAAACCGATCTCGACAAGCTTCCTTTTTTCAAATGCGTCATCAAGGAAACCCTTCGACTCCACCCGCCCATC
301  E  E  T  D  L  D  K  L  P  F  F  K  C  V  I  K  E  T  L  R  L  H  P  P  I
    CCCCTCCTCCTCCACGAAACCGCGGAAGAGACCACCGTAATGGGCTACCGGATACCCGCCCAAGCGCGAGTAATG
326  P  L  L  L  H  E  T  A  E  E  T  T  V  M  G  Y  R  I  P  A  Q  A  R  V  M
    GTCAACTCATATGCCATTGGGAGAGATAAGGACTCGTGGGATGACCCTGATATGTTCAAACCCGAAAGGTTTATG
351  V  N  S  Y  A  I  G  R  D  K  D  S  W  D  D  P  D  M  F  K  P  E  R  F  M
    AAGGAAGGAGTGCCGGATTTCAAAGGAGGGAATTTCGAGTTCCTTCCATTCGGGTCGGGTCGGAGGTCATGCCCG
376  K  E  G  V  P  D  F  K  G  G  N  F  E  F  L  P  F  G  S  G  R  R  S  C  P
    GGAATGCAATTGGGTCTCTATGCGCTTGAGATGACTGTTGCTAATCTGATTCAATGCTTCACGTGGGAGTTACCT
401  G  M  Q  L  G  L  Y  A  L  E  M  T  V  A  N  L  I  Q  C  F  T  W  E  L  P
    AATGGAATGAAGCCAAGCGAGTTGGATATGACTGACATGTTTGGGCTCACTGCACCTCGGGCGAGTCGACTCATT
426  N  G  M  K  P  S  E  L  D  M  T  D  M  F  G  L  T  A  P  R  A  S  R  L  I
    GCTGTGCCAAGACTCAGGCTGACGTGTCCGATCTCTTCATCTTGAGGAAAGAGAAGGAATTGAGATTTCTCCTGC
451  A  V  P  R  L  R  L  T  C  P  I  S  S  S  *
    GGAATGGTATTGCAATGGATGGGGAAAAAAACTAAACAAAAAATCAAAATTAGGGGAGGAAGACTTAAATTTCGT
    CTTCTTTTTTTACCGATTTGCTCGGCTTGATTTTATTTTTTTCCTGTTCATGTGAAAATTGTATTTTTACACATC
    TTTTTCTTCTTTTTTGGTCTTCTTTTTTTAT
```

图 2　*FtF5H* 核苷酸序列及氨基酸序列

Fig. 2　The nucleotide sequence and amino acid sequence of *FtF5H*

2.2 *FtF5H* 基因编码蛋白的生物信息学分析

2.2.1 理化性质分析

利用 ProtParam 软件对 FtF5H 蛋白理化性质进行预测，发现 FtF5H 蛋白分子式为 $C_{2346}H_{3677}N_{637}O_{680}S_{28}$，理论分子量 52 583.54 Da，亲水性平均系数（GRAVY）为-0.234，不稳定系数(II)为 36.57，为稳定蛋白，脂溶指数为 82.65，负电荷残基（Asp+Glu）63 个，正电荷残基（Arg+Lys）56 个，理论等电点 5.76，偏酸性。推测该蛋白为亲水性稳定酸性蛋白。通过 Conserved domains 数据库对 FtF5H 保守结构域进行分析，发现该蛋白有一个 P450 保守域，属于 P450 超家族（图 3）。

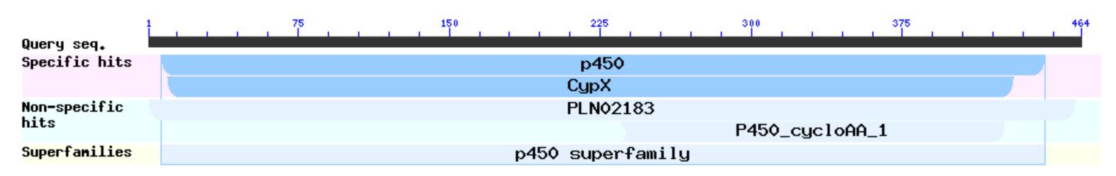


图 3　苦荞 FtF5H 蛋白保守结构域分析

Fig. 3　Analysis of the conserved domains of FtF5H protein

通过 Protscale 在线软件，以 Hphob./Kyte & Dodittle 为标度，对苦荞 FtF5H 蛋白质进行了分析，以氨基酸序列当横坐标，氨基酸标度为纵坐标，亲水性越强，则分值越小。FtF5H 最高分值 1.967 出现在多肽链中的第 256 位氨基酸，该位点的氨基酸疏水性最强；最低分值 -2.856 出现在 362 位氨基酸，证明该位点的氨基酸亲水性最强。NetPhos 3.1Server 预测显示：FtF5H 整条多肽链中分值在阈值 0.5 以上的氨基酸位点共有 36 个，其中丝氨酸（Ser）22 个，苏氨酸（Thr）13 个，酪氨酸（Tyr）1 个。这些位点可能发生磷酸化反应，从而对蛋白的活性与功能产生调控作用。

通过 SignalP 5.0 Server 在线软件对苦荞、苹果（XP 008372753.2）、白梨（AGR44939.1）、喜树（AAT39511.1）、水稻（BAF43423.1）、桃（XP 007203643.1）、柳枝稷（AFH89638.1）、甜樱桃（XP 021802881.1）、藜麦（XP 021718110.1）及大叶藻（KMZ63113.1）10 种植物的 F5H 基因的蛋白质信号肽进行预测比对，结果均不存在信号肽，为非分泌性蛋白（表 2）。因此推测 F5H 基因可能在游离核糖体上合成后不经蛋白转运，直接在细胞质基质的特定部位中行使催化功能（张太奎等，2013）。利用 Psort 对该 10 种植物 F5H 蛋白进行亚细胞定位预测，发现除柳枝稷定位在细胞骨架外，其余均在叶绿体。基于 F5H 基因编码蛋白的信号肽和亚细胞定位进行 SPSS 聚类分析，结果见图 4，双子叶陆生植物桃、甜樱桃、苹果、白梨、苦荞、喜树、藜麦聚为一支，常年生长在水中的水稻、大叶藻聚为一支，柳枝稷单独为一支，推测与 F5H 的作用机制有关。

TMHMM 软件预测显示苦荞、柳枝稷、大叶藻不具有跨膜结构域。苹果、白梨、喜树、水稻、桃、甜樱桃、藜麦均各含有 1 个跨膜结构域，说明植物种类不同，其 F5H 基因编码蛋白质的跨膜区域存在情况不同。

表 2　10 种植物 F5H 基因编码蛋白质信号肽的对比分析

Table 2　Signal peptide prediction for proteins encoded by F5H of 10 plants

物种 Species	氨基酸最高原始剪切位点 Highest original cleavage site of amino acid		最高信号肽剪切位点 Highest signal peptide cleavage site		最高综合剪切位点 Highest comprehensive splicing site		氨基酸残基的平均信号肽 Average signal peptide of amino acid residues		信号肽 Signal peptide
	位点 Site	得分 Score	位点 Site	得分 Score	位点 Site	得分 Score	位点 Site	得分 Score	
苦荞 *Fagopyrum tataricum*	31	0.113	40	0.151	33	0.385	1-39	0.200	NO
苹果 *Malus domestica*	33	0.153	33	0.202	9	0.548	1-32	0.268	NO
白梨 *Pyrus bretschneideri*	33	0.134	14	0.163	9	0.511	1-13	0.268	NO
喜树 *Camptotheca acuminata*	38	0.152	38	0.183	32	0.402	1-37	0.221	NO
水稻 *Oryza sativa*	28	0.360	4	0.256	1	0.631	1-3	0.611	NO
桃 *Prunus persica*	32	0.157	11	0.169	8	0.439	1-10	0.308	NO
柳枝稷 *Panicum virgatum*	56	0.142	56	0.119	54	0.132	1-55	0.098	NO
樱桃李 *Prunus avium*	32	0.157	11	0.163	8	0.415	1-10	0.295	NO
藜麦 *Chenopodium quinoa*	34	0.133	18	0.144	10	0.307	1-17	0.185	NO
大叶藻 *Zostera marina*	22	0.192	2	0.247	1	0.598	1-1	0.598	NO

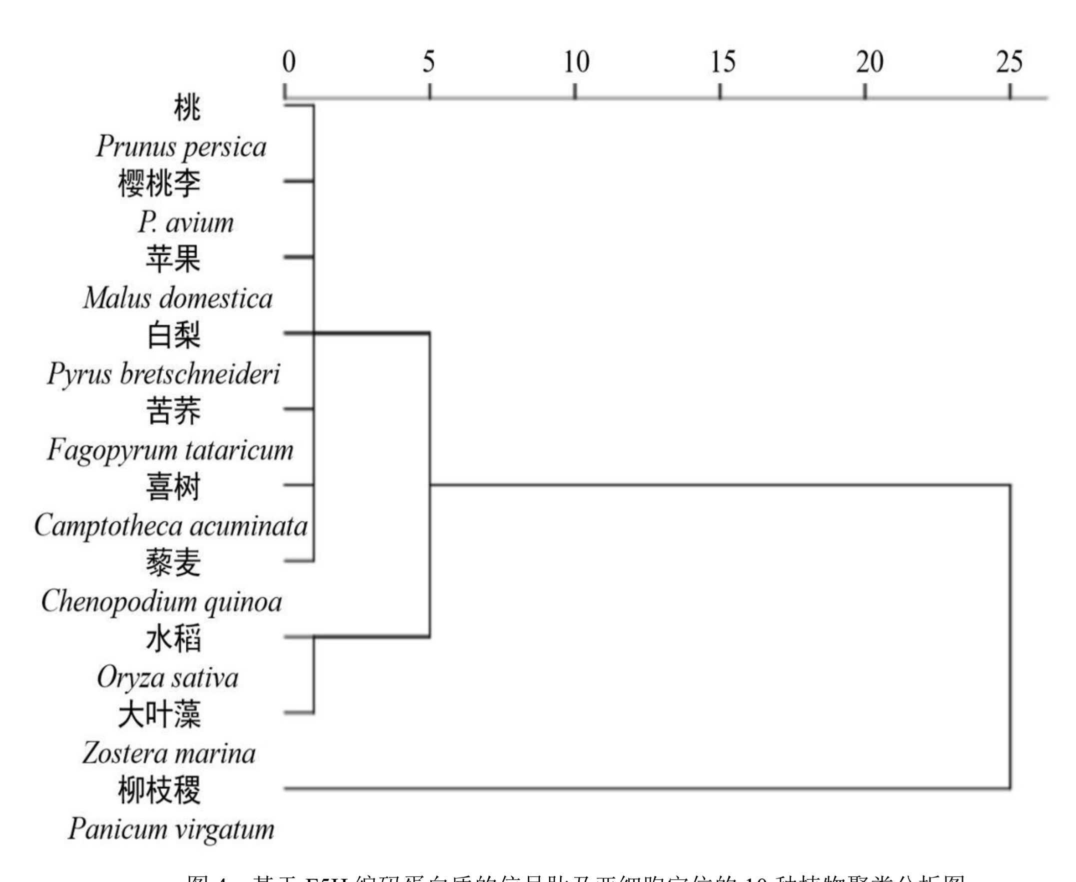
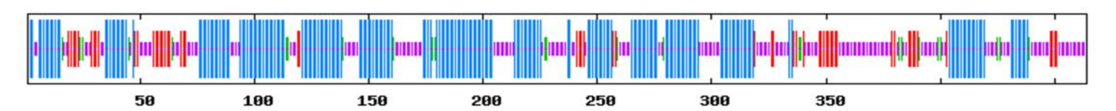


图 4　基于 F5H 编码蛋白质的信号肽及亚细胞定位的 10 种植物聚类分析图

Fig. 4　A cluster diagram based on Signal Peptide and subcellular localization predicted on proteins encoded by F5H of 10 species of plants

2.2.2 FtF5H 的结构预测

通过 NPSA 对蛋白质的二级结构进行在线预测，发现 FtF5H 含有丰富的二级结构，无规则卷曲占 36.21%，α-螺旋占 46.12%，延伸链占 12.28%，β-折叠占 5.39%（图 5）。利用 SWISS-MODEL 对蛋白质的三级结构进行预测，结果见图 6：A。FtF5H 蛋白第 5~455 位氨基酸与数据库目标蛋白 5ylw.1.A 的序列相似度为 33.49%。QMEAN 值是-2.83。GMQE（Global Model Quality Estimation，全球性模型质量估测）值为 0.71，GMQE 值在 0~1 之间，越接近 1，建模质量越好，表明建模可信度高。结合 *Malus domestica*（XP 008372753.2）、*Pyrus bretschneideri*（AGR44939.1）、*P. sinkiangensis*（QHS84908.1）、*Camptotheca acuminata*（AAT39511.1）、*Vitis vinifera*（XP 002272644.1）、*Prunus persica*（XP 007203643.1）、*P. dulcis*（BBH07995.1）、*Medicago sativa*（ABB02161.1）、*Prunus avium*（XP 021802881.1）的蛋白质三维结构骨架进行分析，发现该基因编码的蛋白质具有典型的结构域：含铁原卟啉功能域(图 6：B）和锰离子功能域（图 6：C）（张太奎等，2015）。

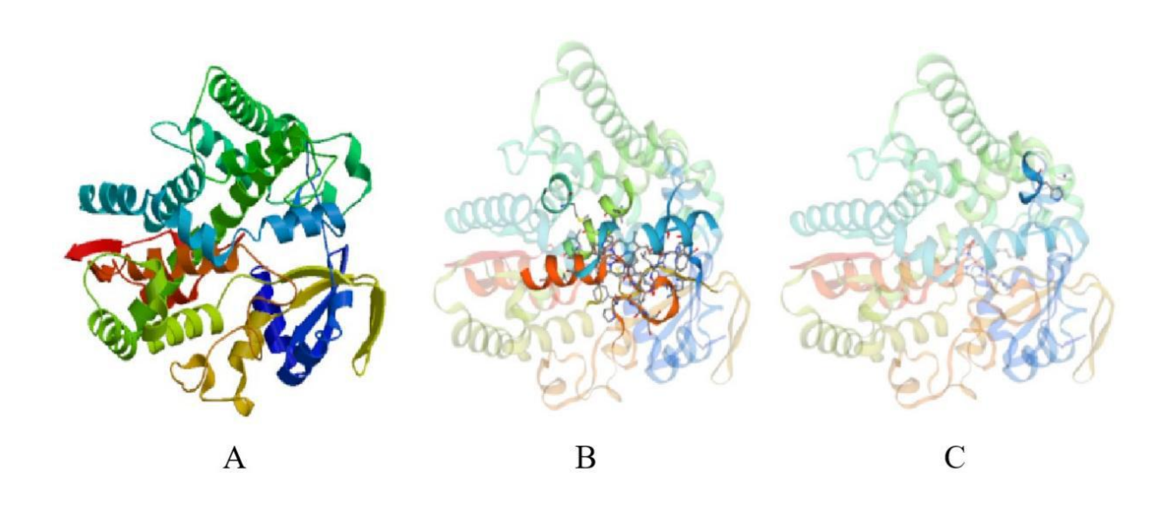


蓝色. α螺旋；红色. 延伸链；绿色. β转角；紫色. 无规则卷曲。

Blue. Alpha helix; Red. Extended chain; Green. Beta turn; Purple. Random curl.

图 5 苦荞 FtF5H 蛋白质二级结构预测

Fig. 5 Prediction of the secondary structure of FtF5H protein



A. FtF5H 蛋白三级结构；B.含铁原卟啉功能域；C.锰离子功能域。

A. Tertiary structure of FtF5H; B. Protoporphyrin ix containing fe; C. Manganese (II) ion.

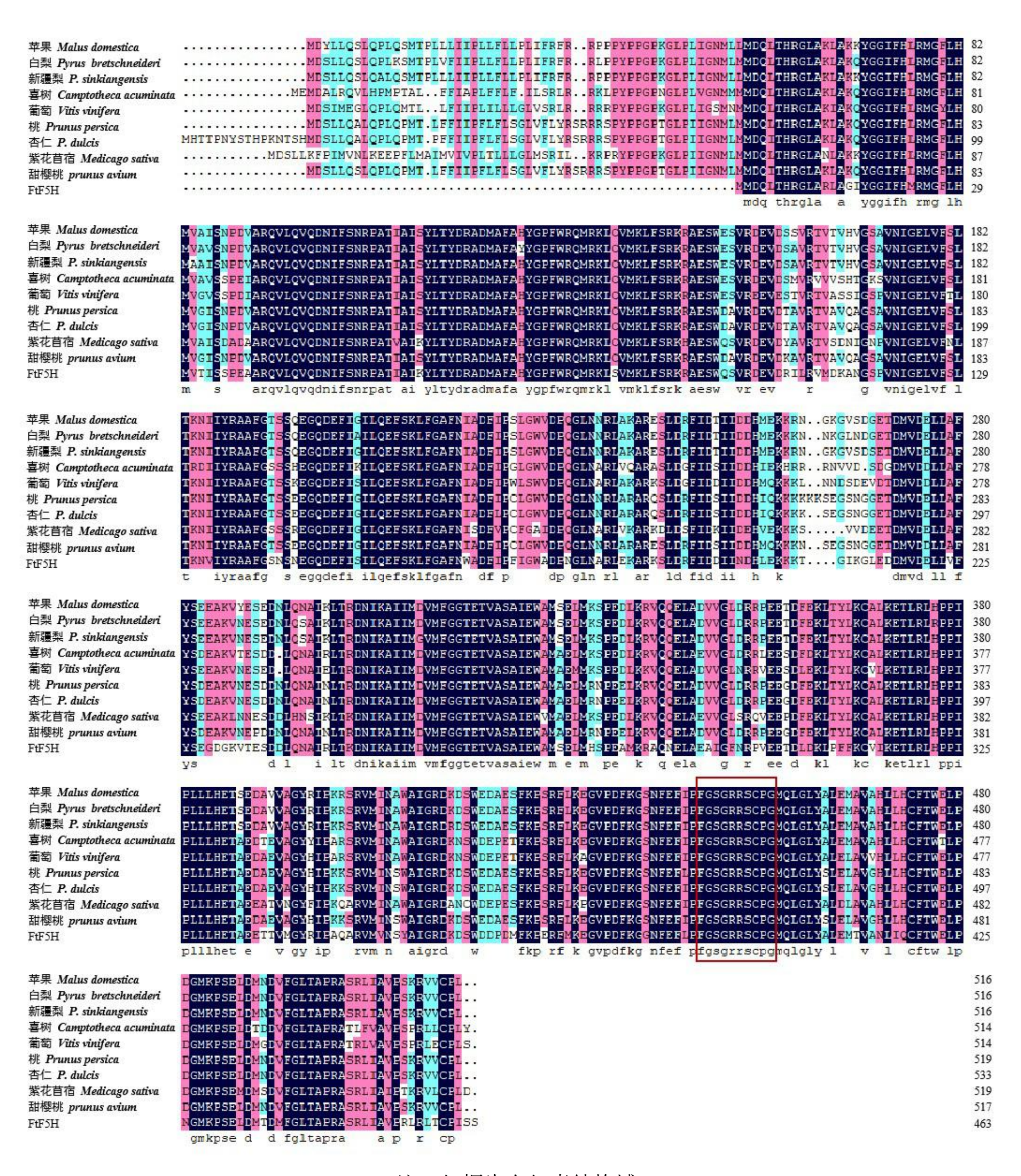
图6 F5H 基因编码蛋白质三级结构及苦荞 FtF5H 蛋白质三级结构预测

Fig. 6 Prediction of three-dimensional structure of protein encoded by F5H gene and three-dimensional structure prediction of FtF5H protein

2.2.3 氨基酸结构分析及系统进化树构建

采用 NCBI 的 BlastP 在线程序，查找 FtF5H 的同源氨基酸序列，比对结果表明 FtF5H 氨基酸序列与 GenBank 中相关序列同源性在 73.12%~77.42%之间。利用 DNAMAN 进行多重序列比对发现，苦荞 FtF5H 与参与比对植物的 F5H 氨基酸序列的相似性高达 84.79%，含有血红素结合结构域 FxxGxxxCxG（图 7），其中保守的半胱氨酸充当血红素铁的第五个配体（Chapple，1998），证明 FtF5H 属于 P450 家族。

运用 MEGA6.0 软件，采用邻接法构建系统进化树，结果显示：苦荞 FtF5H 与藜麦 CYP84A1 聚为一小簇，说明其同源性最高，相关研究报道拟南芥阿魏酸-5-羟化酶 *AtCYP84A1* 的表达与木质素生物合成的定量以及发育调控有关（Ruegger et al.，1999），本研究克隆得到的苦荞 FtF5H 与拟南芥及其它植物的 CYP84A1 聚在一支（图 8），说明其可能与苦荞木质素合成及调控有关（Zhang et al.，2019）。



注：红框为血红素结构域

Note: red frame is heme domain

图 7　苦荞 FtF5H 蛋白的多序列比对

Fig. 7　Multiple sequence alignment of FtF5H protein

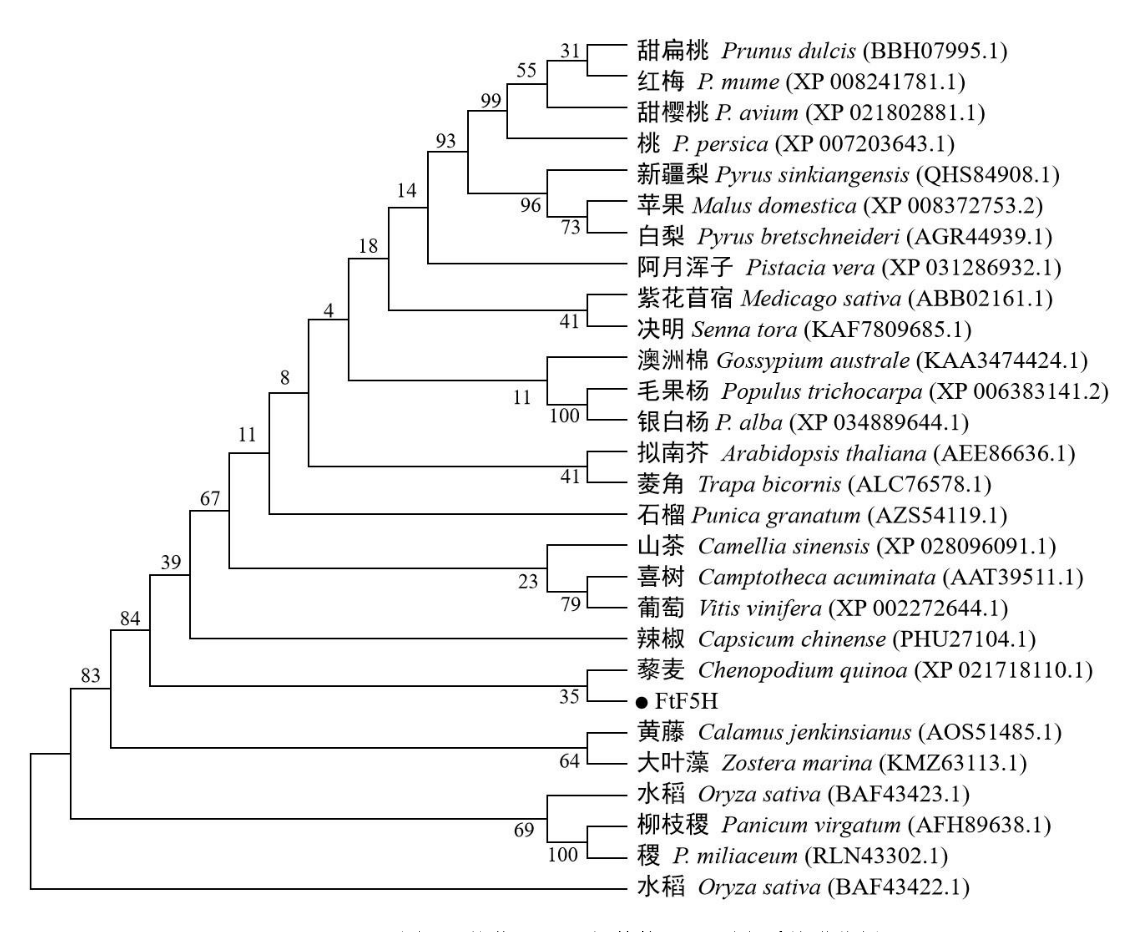


图 8　苦荞 FtF5H 与其他 F5H 蛋白系统进化树

Fig. 8　Phylogenetic tree of FtF5H and other F5H proteins

2.2.4 *FtF5H* 基因相对定量分析

实时荧光定量 PCR（qRT-PCR）分析 *FtF5H* 基因在厚果壳与薄果壳苦荞不同器官的表达水平，结果发现 *FtF5H* 基因在苦荞不同器官的表达量不同，在厚果壳‘云荞 1 号’中的相对表达量由小到大依次为叶、花、茎、果壳；在薄果壳‘小米荞’中则是茎、果壳、花、叶。同一器官相比，除叶外，*FtF5H* 在厚果壳‘云荞 1 号’各个部位的表达量均高于薄果壳‘小米荞’，且差异极显著（图 9）。*FtF5H* 在苦荞厚果壳中的表达量是薄果壳的 5 倍，与前期转录组实验结果基本一致。

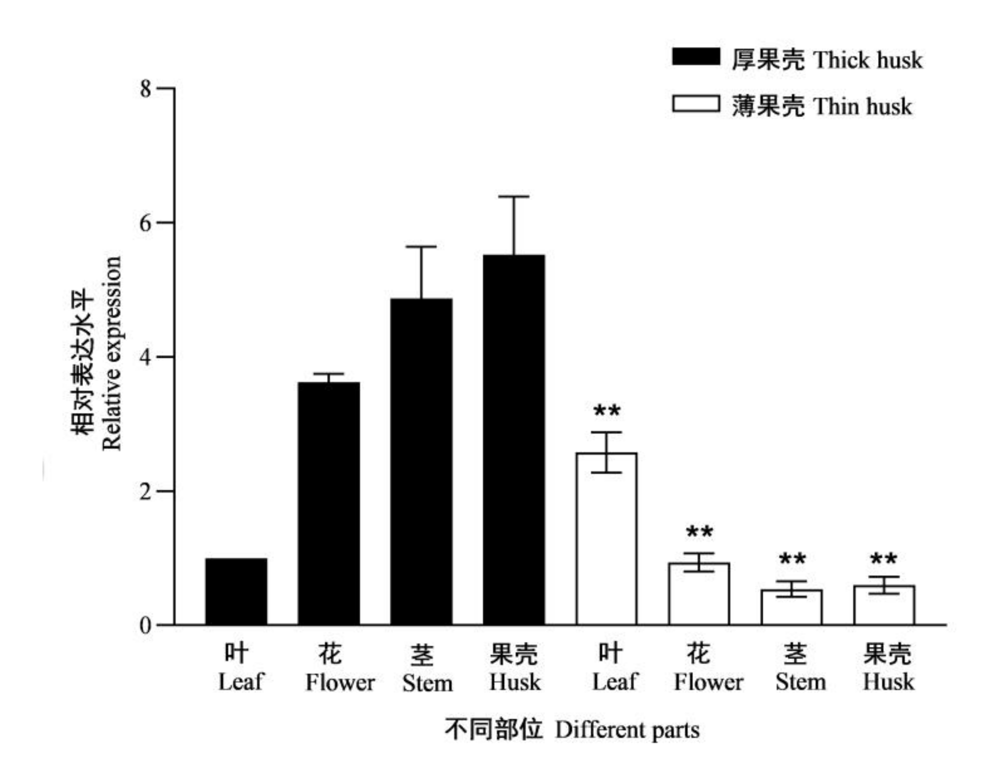


图 9　*FtF5H* 基因在苦荞不同组织部位的相对表达水平

Fig. 9　Relative expression levels of *FtF5H* gene in different tissues of tartary buckwheat

3 讨论与结论

苦荞作为自然界中少见的药食两用作物，具有极高的药用价值及营养功效。如何将薄果壳性状转入栽培品种，育成既易脱壳又高产优质的苦荞新品种成为研究热点。Song et al.（2019）对厚壳苦荞和薄壳苦荞纤维素和木质含量变化与脱皮效率之间关系进行研究，结果表明随着脱壳效率的降低（厚壳），木质素含量减少，纤维素含量增加；吴朝昕（2020）发现薄壳苦荞果壳纤维素和木质素含量显著低于厚壳苦荞，二者研究结果不同。课题组通过转录组测序发现木质素生物合成途径大部分基因在苦荞薄果壳表达量低于厚果壳。为了进一步探究苦荞薄果壳形成的分子机制，本研究克隆了薄果壳和厚果壳的 F5H 基因，序列比对发现二者序列完全一致，且与转录组测序结果一致。

F5H 是调控植物木质素合成的关键酶基因，在木质素形成过程中起重要作用，通过生物信息学分析，发现克隆的苦荞 F5H 蛋白有一个 P450 保守域，属于 P450 超家族，与石榴（冯立娟等，2018）、尾叶桉（肖玉菲等，2018）、毛白杨（王轶男等，2014）中的研究结果一致。苦荞 F5H 蛋白有数量不同的丝氨酸、苏氨酸、酪氨酸 3 个磷酸化位点，与石榴（冯立娟等，2018）中的研究结果一致。F5H 磷酸化后可以改变蛋白质活性，从而调节苦荞的生长发育过程。多重序列比对发现苦荞 FtF5H 与其它植物 F5H 氨基酸序列的相似性高达 84.79%，证明该基因比较保守。系统进化分析结果表明与苦荞 *FtF5H* 基因亲缘关系最近的为藜麦，

且与其它双子叶植物的 F5H 基因亲缘较近，而与单子叶植物水稻、柳枝稷、稷、黄藤的亲缘关系较远，表明该基因在双子叶及单子叶植物纲之间同源性差异较大。

Shafrin et al.（2015）下调黄麻的 F5H 基因，发现与非转基因植物相比，整个茎中酸不溶性木质素含量降低约 25%，纤维木质素降低 12%~15%。肖玉菲等（2018）在对尾叶桉的研究中发现，*EuF5H* 在半木质化茎中表达量最高，在嫩茎中最低。甘蓝型油菜植株根、根茎和茎中木质素含量高低与抗倒能力成正相关，且 F5H 基因在抗倒伏油菜薹期根部、茎部以及开花期根部表达量明显高于易倒伏材料（李尧臣等，2011）。本研究 *FtF5H* 基因在厚果壳‘云荞 1 号’茎、花、果壳的表达量均高于薄果壳‘小米荞’，各器官之间的表达有极显著差异，且 *FtF5H* 在厚果壳中的表达量是薄果壳的 5 倍，推测 *FtF5H* 基因在薄果壳苦荞低表达与薄果壳的木质素合成和积累少有关。

Takeda et al.（2017）对水稻的研究中发现，*OsF5H1* 表达是控制水稻细胞壁中 S/G 木质素组成的主要因素。徐超等（2015）对砀山酥梨果实 F5H 基因的研究发现 *PbF5H* 参与调控梨果实 S 木质素单体的合成，影响木质素 G/S 比值。Tetreault et al.（2020）的研究发现：高粱 F5H（*SbF5H*）的过表达增加了 S-木质素含量。木质素合成有着复杂的代谢网络，多个基因共同调控它的代谢。F5H 基因作为一个多基因家族，并非所有 F5H 基因均在 G/S 的转化中发挥作用，因此本研究获得的 *FtF5H* 基因在薄果壳苦荞木质素生物合成途径的功能仍需进一步研究。

参考文献：

CHAPPLE C, 1998. Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases[J]. Annu Rev Plant Physiol Plant Mol Biol, 49(1) : 311-343.

CHEN QF, CHEN QJ, SHI TX, et al., 2015. Inheritance of tartary buckwheat thick shell character and its Relationships with Yield Factors[J]. Crops, 2 : 27-31. [陈庆富，陈其饺，石桃雄，等，2015. 苦荞厚果壳性状的遗传及其与产量因素的相关性研究[J]. 作物杂志，2 : 27-31.]

CHEN X, CHEN XM, GAI Y, 2015. Lignin biosynthesis in tobacco with regulated expression of poplar ferulate 5-hydroxylase[J].Guangdong Agric Sci, 42(11) : 42-47. [陈雪，陈雪梅，盖颖，2015. 毛白杨阿魏酸-5-羟基化酶对烟草木质素的影响[J]. 广东农业科学，42(11) : 42-47.]

CHEN TY, LIU J, YUAN Y, et al., 2017. Bioinformatics analysis and expressed level of Auxin response factor genes in *Scutellaria baicalensis*[J]. Acta Pharm Sin, 52(11) : 1770-1776.[陈媞颖，刘娟，袁媛，等，2017. 黄

芩 ARF 基因家族生物信息学及表达分析[J]. 药学学报，52(11) : 1770-1776.]

FRANKE R, MCMICHA CM, MEYER K, et al., 2000. Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase[J]. Plant J, 22(3) : 223-234.

FENG LJ,YIN YL,YANG XM, et al., 2018. Identification and Bioinformatics Analysis of *F5H* Gene Family in Pomegranate[C]//Pomegranate Branch of the Chinese Horticultural Society. Progress in Research on Pomegranate in China (3)—The 3rd Representative Conference of the Pomegranate Branch of the Chinese Horticultural Society and the First China Sihong Soft Seed Pomegranate Summit Forum, Proceedings of the National Pomegranate Industry Technology Innovation Alliance Preparatory Meeting. Sihong: Chinese Horticultural Society:9. [冯立娟，尹燕雷，杨雪梅，等，2018.石榴 F5H 基因家族的鉴定与生物信息学分析[C]//中国园艺学会石榴分会.中国石榴研究进展（三）—第三届中国园艺学会石榴分会会员代表大会暨首届中国泗洪软籽石榴高峰论坛、国家石榴产业科技创新联盟筹备会论文集[C].泗洪:中国园艺学会:9.]

HUMPHREYS JM, CHAPPLE C, 2002. Rewriting the lignin roadmap[J]. Curr Opin Plant Biol, 5(3) : 224-229.

LI YC, QI CK, 2011. Lignin content and key gene expression in lignin synthesis of *Brassica napus* L.with lodging resistance[J]. Jiangsu J Agric Sci, 27(3) : 481-487.[李尧臣，戚存扣，2011.抗倒伏甘蓝型油菜(*Brassica napus* L.)木质素含量及木质素合成关键基因的表达[J].江苏农业学报, 27(3) : 481-487.]

LI Y, GU H, QI CQ, 2013. Expression of lignin synthesis key genes of F5H, 4CL and COMT on lodging resistance of *Brassica napus* L[J].Chin J Oil Crop Sci, 35(1) : 24-28. [李扬，顾慧，戚存扣，2013.甘蓝型油菜木质素合成关键基因 F5H、4CL 和 COMT 的定量表达[J]. 中国油料作物学报，35(1) : 24-28.]

MEYER K, CUSUMANO JC, SOMERVILLE C, et al., 1996. Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases[J]. Proc Natl Acad Sci USA, 93(14) : 6869-6874.

MEYER K, SHIRLEY AM, CUSUMANO JC, et al., 1998. Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis[J]. Proc Natl Acad Sci USA, 95 (12) : 6619-6623.

RUEGGER M, MEYER K, CUSUMANO JC, et al., 1999. Regulation of ferulate-5-hydroxylase expression in Arabidopsis in the context of sinapate ester biosynthesis[J]. Plant Physiol, 119(1) : 101-110.

SHAFRIN F, DAS SS, SANAN-MISHRA N, et al., 2015. Artificial miRNA-mediated down-regulation of two monolignoid biosynthetic genes(C3H and F5H) cause reduction in lignin content in jute[J]. Plant Mol Biol, 89(4-5) : 511-527.

SONG C, MA CR, XIANG DB, 2019. Variations in accumulation of lignin and cellulose and metabolic changes in seed hull provide insight into dehulling characteristic of tartary buckwheat seeds[J]. Int J Mol Sci, 20(3) : 524.

TAKEDA Y, KOSHIBA T, TOBIMATSU Y, et al., 2017. Regulation of CONIFERALDEHYDE 5-HYDROXYLASE expression to modulate cell wall lignin structure in rice[J]. Planta, 246(2) : 337–349.

TETREAULT HM, GRIES T, PALMER NA, et al., 2020. Overexpression of ferulate 5-hydroxylase increases syringyl units in *Sorghum bicolor*[J]. Plant Mol Biol, 103(3) : 269–285.

WANG J, 2009.Cloning and Expression Analysis of Flax (*Linum usitatissimum*) Critical Lignin Metabolism Genes[D]. Beijing:Chinese Academy of Agricultural Sciences: 1-66.[王进，2009.亚麻（*Linum usitatissimum*）木质素合成关键酶基因的克隆及表达分析[D]. 北京：中国农业科学院：1-66.]

WANG YN, CHEN X, GAI Y, 2014. Cloning and bioinformatics analysis on Lignin synthetase gene Ferulate-5-hydroxylase of *Populus tomentosa*[J].Guangdong Agric Sci,41(20):131-135. [王轶男，陈雪，盖颖. 2014. 毛白杨木质素合成酶基因 F5H 克隆与生物信息学分析[J].广东农业科学,41(20): 131-135.]

WEN SC, 2015. Effect of Potassium on Related Genes Expression of Ferulic Acid Biosynthesis in Chinese Angelica[D]. Lanzhou: Gansu University of Traditional Chinese Medicine:1-87. [温随超，2015.钾素营养对当归阿魏酸生物合成相关基因表达的影响[D]. 兰州：甘肃中医药大学：1-87.]

WU CX, 2020. Shell structure of rice tartary buckwheat and its transcriptome analysis during development[D]. Guiyang: Guizhou Normal University:1-82.[吴朝昕，2020. 薄壳苦荞果壳结构及其发育中的转录组学分析[D].贵阳：贵州师范大学：1-82.]

XU C, FANG Z, YANG FM, et al., 2015. Analysis of F5H expression and stone cell development in Pyrus bretschneideri cv. Dangshan Su Fruit[J]. Plant Physiol J, 51(5) :778-784.[徐超，方志，杨芳梅，等，2015. 砀山酥梨果实 F5H 表达与石细胞发育的分析[J]. 植物生理学报，51(5) : 778-784.]

XIAO YF, LIU HL, LIU XS, et al., 2018. Cloning, expression and sequence analysis of F5H gene in *Eucalyptus urophylla* clone GLU4[J]. Bio, 28(3) : 205-211.[肖玉菲，刘海龙，刘雄盛，等，2018. 尾叶桉 GLU4 无性系 F5H 基因的克隆表达及序列分析[J].生物技术，28(3) : 205-211.]

ZHANG TK, LIU Z, ZHU FM, et al., 2013. Bioinformatics analysis of chalcone synzyme genes from ten kinds of ornamental plants[J]. J West Chin For Sci, 42(5) : 62-68.[张太奎，刘峥，朱芳明，等， 2013. 10 种观赏植物查尔酮合成酶基因生物信息学分析[J].西部林业科学, 42(5) : 62-68.]

ZHU YH, GUO YX, 2014. Research progress in development and utilization of tartary buckwheat resource in

China[J].Sci Technol Food Ind, 35(24) : 360-365. [朱云辉, 郭元新, 2014. 我国苦荞资源的开发利用研究进展[J].食品工业科技, 35(24) : 360-365.]

ZHANG TK, ZONG D, ZHU FM, et al., 2015. The marker gene ACO of fruits preservation[J]. J Yuannan Agric Univ(Nat Sci Ed), 30(2) : 245-251.[张太奎, 纵丹, 朱芳明, 等, 2015. 果树果实保鲜力强弱标记基因ACO[J]. 云南农业大学学报(自然科学), 30(2) : 245-251.]

ZHANG TK, LIU CY, HUANG XB, et al., 2019. Land-plant phylogenomic and pomegranate transcriptomic analyses reveal an evolutionary scenario of *CYP*75 genes subsequent to whole genome duplications[J]. J Plant Biol , 62 : 48-60.